



RICE | HERC
Houston Education Research Consortium

Building Better Cities
Building Better Lives

Rice University's Kinder Institute for Urban Research



Investigating Causal Effects of Arts Education Experiences:

Experimental Evidence from
Houston's Arts Access Initiative

By Daniel H. Bowen, Ph.D. and Brian Kisida, Ph.D.

Phil Roeder

Research Report
for the Houston Independent School District

Volume No. 7, Issue No. 4 | February, 2019

Research Report

Investigating Causal Effects of Arts Education Experiences: Experimental Evidence from Houston's Arts Access Initiative

Daniel H. Bowen, Ph.D.¹

Assistant Professor

Department of Educational Administration &
Human Resource Development
Texas A&M University
dhbowen@tamu.edu

Brian Kisida, Ph.D.

Assistant Professor

Truman School of Public Affairs
University of Missouri
kisidab@missouri.edu

¹ Corresponding author; email: dhbowen@tamu.edu

Abstract

The recent wave of test-based accountability reforms has negatively impacted the provision of K-12 arts educational experiences. Advocates contend that, in addition to providing intrinsic benefits, the arts can positively influence academic and social development. However, the empirical evidence to support such claims is limited. We conducted a randomized controlled trial with 10,548 3rd-8th grade students who were enrolled in 42 schools that were assigned by lottery to receive substantial influxes of arts education experiences provided through school-community partnerships with local arts organizations, cultural institutions, and teaching-artists. We find that these increases in arts educational experiences significantly reduce the proportion of students receiving disciplinary infractions by 3.6 percentage points, improve STAAR writing achievement by 0.13 of a standard deviation, and increase students' compassion for others by 0.08 of a standard deviation. For students in elementary schools, which comprise 86 percent of the sample, we find that these arts educational experiences also significantly improve school engagement, college aspirations, and arts-facilitated empathy. These findings provide strong evidence that arts educational experiences can produce significant positive impacts on student academic and social development. Policymakers should consider these multifaceted educational benefits when assessing the role and value of the arts in K-12 schools.

Acknowledgments

We are grateful for the guidance, access, and support we received from the Houston Independent School District, the participating evaluation schools, the Houston Education Research Consortium, Young Audiences Inc. of Houston, and the participating arts partners and their administrators and staff. We also thank the following individuals who were instrumental, helpful, and supportive of our efforts: Daniel Gohl, Dr. Andrew Houlihan, Dr. Grenita Lathan, Deborah Lugo, Wenden Sanders, Adam Stephens, Carla Stevens, Douglas Torres-Edwards, and all of the school principals, campus arts liaisons, and administrators and representatives from arts organiza-

tions who participated in our focus groups, interviews, and surveys. Finally, we acknowledge the incredible research assistance and feedback we received from Dr. Holly Heard, Brenna Lin, Dr. Jean Madsen, Kelsey Ragan, Dr. Ruth N. López Turley, and the panel discussants and participants at our Arts Education Partnership, Association for Education Finance and Policy, Education Policy Collaborative (EPC), and University Council for Educational Administration conference presentations. This study was made possible by generous support of the Houston Endowment, National Endowment for the Arts, and the Spencer Foundation.

Arts Access Initiative Arts Organizations for 2015–16 and 2016–17

AFA	Da Camera	Houston Grand Opera	Musiq
Alley Theatre	Ensemble Theatre	Houston Symphony	Nameless Sound
Aperio, Music of the Americas	Express Children's Theatre	Houston Youth Symphony	Prelude Music Foundation
Ars Lyrica	FotoFest Inc.	Main Street Theater	Society for the Performing Arts
Arte Público Press	The Hobby Center for the Performing Arts	MECA	Texan-French Alliance for the Arts
Art League Houston	Houston Ballet	Mercury Chamber Orchestra	Theatre Under the Stars
City ArtWorks	Houston Center for Photography	METdance	Writers in the Schools
Contemporary Arts Museum Houston	Houston Chamber Choir	Museum of Fine Arts, Houston	Young Audiences of Houston

Introduction

Following a steady increase throughout most of the 20th century, arts education opportunities in the United States have been in steady decline since the 1980s (Hamilton et al., 2007; Rabkin & Hedberg, 2011; von Zastrow & Janc, 2004; West, 2007). This trend has not been equitable across student subgroups. Over this period, white students have experienced virtually no change, whereas African-American students have experienced reductions of 49 percent and Hispanic/Latinx students by 40 percent. Moreover, children whose parents have less than a high school education have experienced a 77 percent decline (Rabkin & Hedberg, 2011). Childhood arts experiences predict arts participation and engagement later in life, meaning that decreases in childhood access likely has lasting effects. These trends are especially concerning in Houston, where 86% of students enrolled in the Houston Independent School District (HISD) are African-American or Hispanic/Latinx (HISD, 2017). Moreover, in the 2012 Houston Arts Survey, it was reported that nearly 60% of respondents had attended at least one artistic performance in the past year; however, participation plummeted to 29% for individuals from households with incomes at or below \$37,500 (Klineberg, Wu, & Aldape, 2012).

The arts have intrinsic benefits for participants, but advocates also contend that the arts play a vital role in public education because they enhance skills and knowledge that transfer to student performance in other academic subjects, positively affect social and emotional learning, and enhance artistic ability and creativity which are valuable skills in today's economy. (Deasy, 2002; Eisner, 2002; Winner, Goldstein, & Vincent-Lancrin, 2013; Winner & Hetland, 2001). However, the arts remain a prime target for cuts when education administrators and policymakers face accountability-driven tradeoffs (Yee, 2014). Since the passing of the No Child Left Behind Act (2001), the emphasis on standardized testing in "core subjects" has

coincided with notable declines in school-facilitated arts exposure (Gadsden, 2008). There is a strong link between the emphasis on accountability testing and decreases in time and resources for the arts and other non-tested subjects (Bassok, Latham, & Rorem, 2016; Government Accountability Office, 2009; Guzenhauser, 2003; West, 2007). Moreover, evaluations of education programs and policies increasingly rely on empirical evidence, yet the causal benefits of the arts have rarely been rigorously investigated with experimental methods (Winner & Cooper, 2000; Winner, Goldstein, & Vincent-Lancrin, 2013; Winner & Hetland, 2001). As a result, policymakers and administrators struggle to make the case for the arts



School participation in the Initiative was voluntary, and applicants were required to commit to a monetary match earmarked for arts experiences through teaching-artist residencies, in-school professional artist performances, field trips, and afterschool programs. Demand for AAI participation exceeded supply in the first two years of implementation. Forty-six campuses applied to participate in the first year of the AAI, and 35 additional campuses applied to participate in the second year. After consulting with our research team, AAI stakeholders agreed to randomly allocate participation among 42 schools in the first two years and

in K-12 schools, and advocates lack evidence to quantify the costs that come with decreases in access.

An increasingly common strategy for addressing K-12 arts educational inequities is the formation of school-community arts partnerships. Typically, these partnerships consist of school- and district-level administrators, cultural institutions, philanthropists, government officials, researchers, and a “backbone” organization that facilitates these collaborations (Bowen & Kisida, 2017; Perille, 2016). Successful partnerships have flourished in many urban school districts, including Boston, Chicago, Los Angeles, and Seattle. One such partnership, Houston’s Arts Access Initiative (AAI), launched in 2013. This multi-sector, collaborative effort aimed to advance student access to the arts through “strategy, partnerships, data collection, and advocacy.” The AAI vision statement was “that every student in Houston will have the opportunity to benefit cognitively, creatively, emotionally, and academically through the arts,” with an emphasis on foundational goals of equity, impact, and sustainability. As with arts education research in general, only a few studies have investigated the effects of these partnerships (Catterall & Waldorf, 1999; Rowe, Werber, Kaganoff, & Robyn, 2004). Preliminary research suggests that school-community arts partnerships improve academic achievement and attainment (Catterall & Waldorf, 1999); however, there is no causal evidence about the impacts of these interventions.

After conducting a district-wide campus inventory, AAI stakeholders developed strategies to serve schools with the lowest levels of arts resources, primarily through forging and enhancing school-community partnerships.

deferring AAI participation for the other applicants. In addition to being an impartial method for selecting AAI schools, this process was conducive to conducting a cluster randomized controlled trial (RCT).

In this study, we investigate the causal effects of increases in elementary and middle school students’ arts educational experiences, provided through school-community arts partnerships in the first year of program implementation. Comparing AAI school gains to their randomly assigned counterparts, we find that increasing arts learning opportunities significantly reduces the proportion of students receiving disciplinary infractions, increases writing achievement, and increases compassion for others. We do not find significant effects overall on student attendance, math, reading, or science achievement, or other survey-based measures. For students in elementary schools, which were the main emphasis of the Initiative and comprised 86 percent of our school sample, we find that the intervention also improves student engagement, college aspirations, and arts-facilitated empathy. We also find that these effects varied substantially by student subgroups. Students receiving “limited English proficiency” (LEP) and gifted and talented (GT) program services appear to benefit more from these arts learning experiences. These results provide strong evidence that substantial increases in arts access can produce significant, multifaceted educational benefits for students.

Previous Research

The benefits of arts education are rich in theory and testimony but remain scant on rigorous empirical evidence. While existing evidence suggests that there are positive, significant associations between arts participation/engagement and educational outcomes, such studies are typically correlational and subject to omitted variable bias (Elpus, 2013). In a systematic review of the literature, Winner et al. (2013) concluded that there is limited causal evidence demonstrating that the arts affect academic outcomes. Notable exceptions include drama education, which appears to strengthen verbal skills (Podlozny, 2000), and music education, which seems to strengthen IQ (Moreno et al., 2011; Neville, 2008; Schellenberg, 2004) and improve math and reading skills (Gromko, 2005; Lee & Kim, 2006; Moreno et al., 2009; Standley, 2008).

The lack of rigorous empirical evidence is attributed to a paucity of experimental investigations. There is also concern that researchers have investigated arts impacts with a limited set of outcomes. Arts education proponents contend that arts learning experiences develop student self-expression, creativity, and empathy (Dewey, 1919; Ruppert, 2006; Zimmerman, 2009), and serve as ways to enhance cognitive abilities which foster critical thinking skills (Eisner, 2002). However, researchers rarely incorporate such outcomes in their investigations. A notable exception was an RCT evaluation of the Crystal Bridges Museum of American Art's school visit program, where researchers collected original data and found that arts exposure increased student tolerance, historical empathy, and the ability to think critically about works of art (Bowen, Greene, & Kisida, 2014; Greene, Kisida, & Bowen, 2014; Kisida, Bowen, & Greene, 2016). In a recent review of arts integration research, this study was the only example of "strong evidence," as defined by the U.S. Department of Education, to demonstrate causal impacts from school-related arts education interventions (Ludwig, Boyle, & Lindsay, 2017). Moreover, a series of experimental

studies of live theater field trip performances provide evidence to suggest that these experiences increase students' tolerance, social perspective taking, and command of a play's plot and vocabulary (Greene, Erickson, Watson, & Beck, 2018). However, the Crystal Bridges and theater performance studies have limited external validity, as these interventions were confined solely to a single experience with one arts organization's field trip program that occurred over a relatively short time span.

Lingering policy-relevant questions also remain regarding who benefits the most from receiving school-sponsored arts education activities. Several studies have identified that race/ethnicity and socioeconomic status are strong predictors of access to arts and cultural experiences outside of school, with historically-underserved populations receiving substantially less exposure through out-of-school, family-facilitated experiences (Kisida, Greene, & Bowen, 2014; Meyer, Princiotta, & Lanahan, 2004; Redford, Burns, & Hall, 2018). Moreover, studies have shown that historically-underserved students demonstrate more-pronounced, positive impacts from school-sponsored arts exposure, suggesting that these interventions reduce gaps in educational outcomes tied to arts-educational activities (Catterall, Dumais, & Hampden-Thompson, 2012; Kinney & Forsythe, 2004; Kisida, Greene, & Bowen, 2014; Podlozny, 2000; Thomas, Singh, & Klopfenstein, 2015). Consequently, students from underserved communities are more dependent on schools to provide arts access to attain these benefits.

School-Community Partnerships

School-community partnerships have become an increasingly popular strategy for addressing inequities in arts education (Bowen & Kisida, 2017). These partnerships tend to be broad coalitions formed between school administrators, cultural institutions and organizations, policymakers, and philanthropists (Perille, 2016). No study to date has examined the causal impacts of these school-community arts partnerships, but there have been noteworthy non-experimental evaluations. The earliest evidence on these partnerships comes from investigations of the school-community arts partnership prototypes that originated in Chicago and Los Angeles in 1992 and 1999, respectively. The Chicago Arts Partnerships in Education (CAPE) emphasized arts integration, where teaching-artists participated in elementary and high schools for in-depth six-week residencies. The CAPE evaluation team concluded that this initiative produced a mix of positive, though mostly null, impacts on student standardized test scores (Catterall & Waldorf, 1999). The Los Angeles Arts for All program was a ten-year endeavor that emphasized the cultivation of school-community arts partnerships as a means for enhancing arts education throughout the Los Angeles Unified School District. The evaluation of this program was strictly qualitative. While this study did not address student outcomes, the researchers concluded that there were inherent problems in the formation of school-community partnerships. They found that active communication, shared goals, buy-in, and adequate resources were essential for partnerships to succeed, but difficult to

maintain. Specifically, arts organizations were often more interested in objectives tied to increasing student exposure to the arts and bolstering positive public relations through “simple transactions,” such as one-off field trip experiences, and less interested in providing resources that were tied to teachers’ educational goals, such as professional development sessions and experiences aligned with curriculum and standards (Rowe et al., 2004).

The Chicago and Los Angeles efforts served as models and provided lessons for a new crop of school-community partnerships. In addition to Houston’s Arts Access Initiative, these partnerships, as well as revised or newer editions of preexisting partnerships, have developed in major metropolitan areas such as Boston (Arts Expansion Initiative), Chicago (Creative Schools Initiative), Dallas (Learning Partners), New Orleans (KID smART), and Seattle (Creative Advantage). While there have not been impact evaluations of the causal impacts of these arts partnership initiatives on students’ educational outcomes, descriptive evaluations suggest that these partnerships are effective mechanisms for improving student access to arts educational resources and experiences and address-

ing deficits and inequities. Studies have found that these partnerships have successfully increased public school arts education funding (Perille, 2016; Silk, 2016); boosted the number of schools working with community arts partners (Silk, 2016); and expanded the number of credit-bearing arts education courses available to students (Gibson, 2016).



Ed Uthman

The AAI Intervention

A critical early step in the AAI was the collection of descriptive data on campus-level arts educational resources throughout HISD. These data were used to construct an inventory for assessing levels and distributions of arts resources and included the number of certified arts specialists, whether a school offered before/after-school arts programs, whether the campus had facilities for providing arts learning opportunities, and the number of partnerships with arts organizations over the course of the prior school year. Some of the key findings from the initial campus inventory revealed that 29 percent of the 209 K-8 campuses had no full-time arts specialist; 39 percent had either one or no community arts partners;² and 30 percent did not provide any arts programming outside of regular school hours (Young Audiences, 2014). While a substantial proportion of HISD's campuses were struggling to provide arts learning opportunities, it did not appear to be due to a lack of support from school leaders. Ninety-eight percent of surveyed principals and teachers agreed that “students benefit from access to the arts in school.”

The findings from the campus inventory motivated AAI stakeholders to address gaps in campus-level arts resources, primarily through funding efforts to develop and enhance partnerships throughout Houston. AAI leadership recruited HISD elementary and middle schools to participate in the inaugural 2015–16 academic year, with an emphasis on serving campuses with the lowest levels of arts resources.³ School participation in the

Initiative was voluntary, and to be considered principals had to commit to the mission of the Initiative, engage in strategic arts planning with the AAI director, designate a campus-level arts liaison to coordinate and facilitate AAI-related efforts, participate in teacher and principal arts-integration professional development, and attend AAI peer-network mentoring sessions. Participating schools were also required to commit a monetary match between \$1 to \$10 for AAI funding, earmarked to provide arts experiences through teaching-artist residencies, in-school professional artist performances, field trips, and afterschool programs.⁴ The AAI director and staff worked with principals to understand their schools' goals for the upcoming year to help guide arts program selections that would align with their objectives. As part of this process, the director and staff encouraged principals to budget for a diversity of programs such that all of the major arts disciplines were included: dance, music, theater, and visual arts.⁵ The Houston Endowment provided a 1:1 match for each school's financial commitment. Including matched funds, AAI schools had an average annual budget of \$14.67 per student to facilitate and enhance partnerships with arts organizations and institutions.

Using the baseline arts campus inventory, we can assess the impact of the AAI on the number of school-community arts partnerships formed as a result of the Initiative. In the year prior to applying to join the AAI, treatment and control schools were statistically equivalent in terms of their average number of school-community partnerships. In the year prior to the Initiative, AAI treatment

2 The term “partnership” is defined here as having worked with or participated in an educational program with a community arts organization or institution. This definition broadly incorporates everything from a one-off field trip to in-depth, longer-term experiences, such as teaching-artist residencies.

3 AAI stakeholders decided to put an initial emphasis on serving elementary, along with a small portion of middle schools in its early phases, with the intention of eventually expanding to serve more secondary schools.

4 Afterschool programs were initially considered to be one of the emphasized arts educational program formats, but this requirement was deemphasized in the initial two years of implementation.

5 There were also organizations and institutions that provided creative writing and hybrid (e.g., musical theater) programs. There were not enough creative writing program providers to make this discipline one that schools were required to include, but these programs were available to be included in AAI schools' plans.



Ed Uthman

schools had 2.76 arts partnerships compared to 2.80 for the control group. Each AAI school averaged 9.86 partnership-facilitated arts educational experiences over the school year, which represented a net increase of 5.03 additional partnerships relative to the control group schools over the same period. Thirty-nine percent of these experiences were provided in-kind by AAI partnering cultural organizations and institutions. Therefore, it is worth noting that the aforementioned average per-student budget underestimates the total investment of resources that students received as a result of the AAI.

The AAI director and staff worked with principals to understand their schools' goals for the upcoming year to help guide arts program selections that would align with their schools' objectives. The principal had the final say on their arts plan, but the AAI director and staff encouraged administrators to select a diverse array of programs such that each of the arts disciplines and formats were provided at least once over the course of the year. Fifty-four percent of AAI student experiences were primarily theatre-based, 12 percent dance, 18 percent music, and 16

percent visual arts;⁶ 31 percent of these student experiences were provided through on-campus professional artist performances, 27 percent were field trip experiences, 33 percent were teaching-artist residencies, and 9 percent were programs provided outside of regular school hours.

⁶ One AAI school participated in a creative writing program.

Study Design

In collaboration with AAI stakeholders, we developed the following research questions to assess the Initiative's impacts on student outcomes:

- Does a substantial increase in arts education experiences improve student-school engagement?
- Do these experiences increase students' desires to engage and participate in the arts?
- Do these experiences increase academic achievement as reflected in standardized test score growth?
- Does the AAI affect students' attitudes and values, specifically in the forms of tolerance, empathy, and compassion for others?
- Are there heterogeneous effects in outcomes across student subgroups?

The outcome measures for this study come from HISD administrative records and original student survey data. HISD records provided us with a dichotomous indicator

for whether a student received a disciplinary infraction over the course of the school year, number of absences, and State of Texas Assessments of Academic Readiness (STAAR) standardized achievement gains in reading, math, science, and writing. STAAR reading and math assessments are administered to all 3rd-8th grade students. The writing assessment is administered to 4th and 7th grade students. The science test is administered to 5th and 8th grade students. We developed our survey constructs using preexisting, established instruments. Additional survey items were developed in coordination with HISD and AAI program providers and stakeholders. These constructs are student-school engagement, college aspiration, arts-facilitated empathy, compassion for others, tolerance, desire to participate in cultural consumption, disposition for arts transfer, and perceived value of the arts. Student survey outcomes, measures of construct reliability, item sources, and individual items are provided in table 1.

With demand for AAI program participation exceeding supply, we were able to conduct a school-level, cluster

RCT. Due to the limited sample size of 42 campuses participating in the AAI evaluation, school applicants were stratified by pairs prior to randomization, based on (in order) grade levels served, student demographics, arts resources, and school-level achievement. This pre-randomization stratification strategy improves statistical precision when estimating effects with a limited number of clusters (Gerber & Green, 2012).

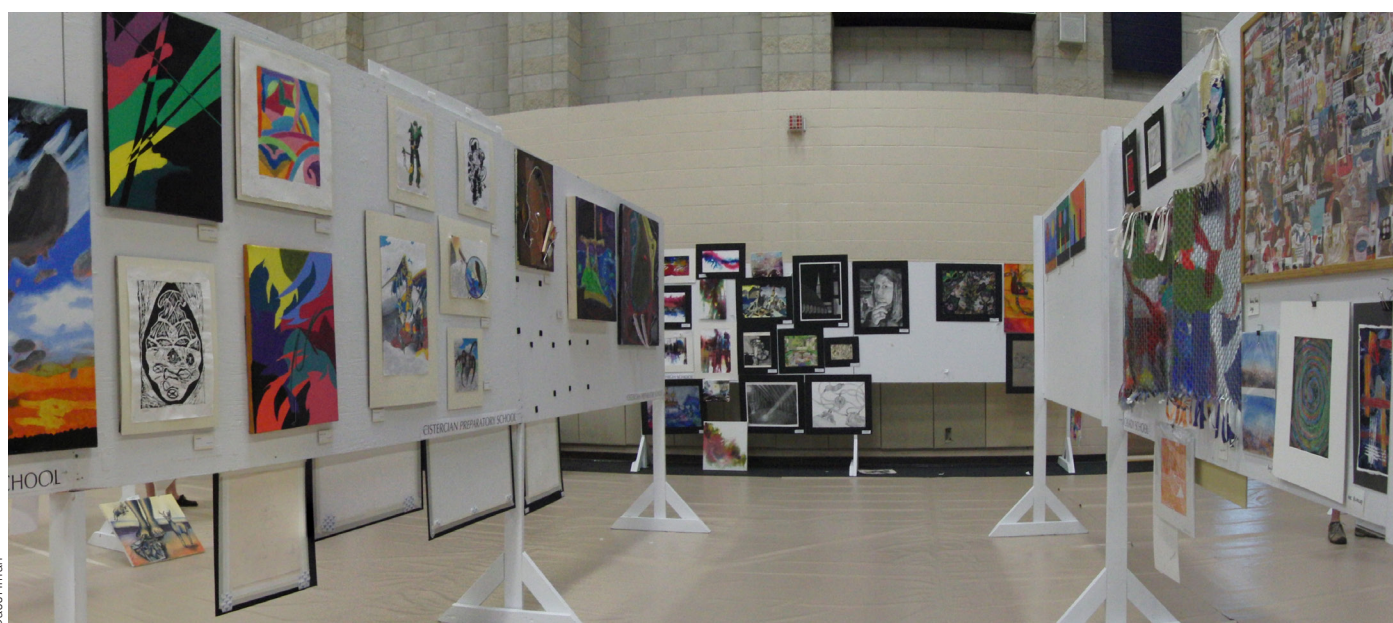


fossimike

Data Collection

Student-level administrative data were obtained from HISD records through the Houston Education Research Consortium's (HERC) longitudinal database, which provides student demographics, school attendance and enrollment records, STAAR scores, and disciplinary records. We collaborated with HISD to collect original survey data in 2016–17, which was year one of implementation for cohort two schools and year two of implementation for cohort one schools. Using student identification codes during survey administration, survey data were linked to student-level administrative records. The grade levels for survey participation were restricted to students enrolled in STAAR-tested grades (3rd–8th). Moreover, at the request of HISD, we restricted our survey population to students who had not received special accommodations when taking the STAAR assessments. Campus testing coordinators administered a pilot survey at the beginning of the fall semester (late September through early October), and an outcome survey at the end of the school year (late April through May). Students' responses on the spring

survey provide outcome measures for both of the AAI evaluation cohorts. We are able to use the piloted fall survey as a measure of baseline responses with cohort two students as a robustness check of our findings. We could not use fall survey responses as a baseline with the first cohort because these students had already participated in the Initiative for a year prior to this round of survey data collection. Moreover, with this being the second year of implementation with the first cohort of AAI schools, we were not able to survey students who graded and transferred out of AAI treatment and control schools after the 2016–17 school year.



DaJHman

Sample

The analytical sample for this study is restricted to 4th-8th grade students with baseline standardized math and reading test scores. We control for baseline test-score achievement throughout our analyses; therefore, students enrolled in 3rd grade and below are not included in these analyses. There were 10,548 students enrolled in AAI treatment and control schools over the evaluation period. There were 1,934 first cohort students who graded and transferred from their AAI evaluation schools. Of the 8,614 AAI evaluation students enrolled in AAI treatment and control schools the year of the survey, we collected and linked survey to administrative data for 6,340 (74 percent) students. Of the 6,500 second cohort AAI evaluation students, we collected and linked 4,901 (75 percent) spring surveys to administrative records; of these second cohort spring survey participants, we were able to link 3,708 (76 percent) of their spring responses to the pilot fall survey (or 57 percent of the second cohort of students). Thirty-six of 42 schools served students at the elementary level, with students in grades PK–5, and six were middle schools with students in grades 6–8. Student demographics, as well as baseline measures of achievement, attendance, and discipline are provided in table 2.

AAI treatment and control school students were statistically equivalent in terms of baseline grade level, gender, free- and reduced-price lunch (FRL) eligibility and higher poverty status, race/ethnicity, proportion of students receiving special education services, proportion of students receiving LEP program services, math achievement, and the proportion of students who had received a disciplinary infraction the prior academic year. However, there was a statistically significant difference in average baseline reading achievement. AAI treatment students had higher STAAR reading test scores than control students in the year prior to participating in the AAI. While this baseline imbalance calls into question the integrity of the randomization process, the fact that there is one signifi-

cant difference across fifteen baseline characteristics can be attributable to a chance outcome due to random assignment. In our analyses, we control for baseline reading achievement, along with the other specified demographic covariates, to improve precision and strengthen our ability to generate unbiased estimates of the treatment on our outcomes of interest (Gerber & Green, 2012).

Comparing targeted student participants to nonparticipants shows the extent to which survey respondents are representative of the targeted population. Survey participants, relative to the targeted survey population, were significantly more likely to be female and white, less likely to have received a disciplinary infraction in the prior academic year, had fewer absences, were less likely to be receiving special education services (due to the restriction of not surveying students who receive accommodations for STAAR testing), and were higher achieving in terms of standardized baseline STAAR reading and math scores. However, when we compare the survey participants by treatment and control status, the only observable statistically significant difference is baseline reading achievement, and this difference is not statistically significantly different from that of the broader sample ($p = 0.67$). Survey participant demographics, by survey participation and treatment status, are available in the appendices.

Analysis

The experimental design of this evaluation provides a straightforward analytical strategy. We estimate the AAI intent-to-treat (ITT) effects with the following model:

$$Y_{ism} = \alpha + AAI_s\gamma_1 + Match_m\gamma_2 + X_i\gamma_3 + \varepsilon_{ism} \quad (1)$$

Where Y signifies an outcome of interest for student, i , enrolled in school-grade, s , matched pair, m . AAI is a dichotomous variable that indicates whether the student was enrolled in a school that was randomly assigned to participate in the AAI; $Match$ is a vector of dummy variables for the pre-randomization matched pairs; X is a vector of student demographics, specifically student grade level, gender, race/ethnicity, FRL eligibility/poverty status, a dichotomous indicator for receipt of special education services, LEP status, prior year absences, an indicator for whether the student received a disciplinary infraction the previous school year, and prior year's STAAR math and reading scores; ε is the error term clustered at the school-grade level. We have also estimated local average treatment effects with a two-staged least squares regression, where the AAI lottery serves as an instrument to predict the percentage of a student's school year spent in an AAI school to then estimate the impact of being enrolled for an entire academic year. First stage estimates, depending on outcome measure, range from 0.961 to 0.987, meaning that local average treatment effects were only slightly larger in magnitude than the ITT estimates. Moreover, none of the local average treatment effects were substantially different than the ITT estimates; these results are provided in the appendices.

We also investigate student-level treatment effect moderators. With such a diverse student population, we examine whether the AAI treatment produces effects that vary by subgroups that have been historically correlated with differences in educational outcomes. Studies have found that effects with arts education interventions tend to be more pronounced with students from historically-under-

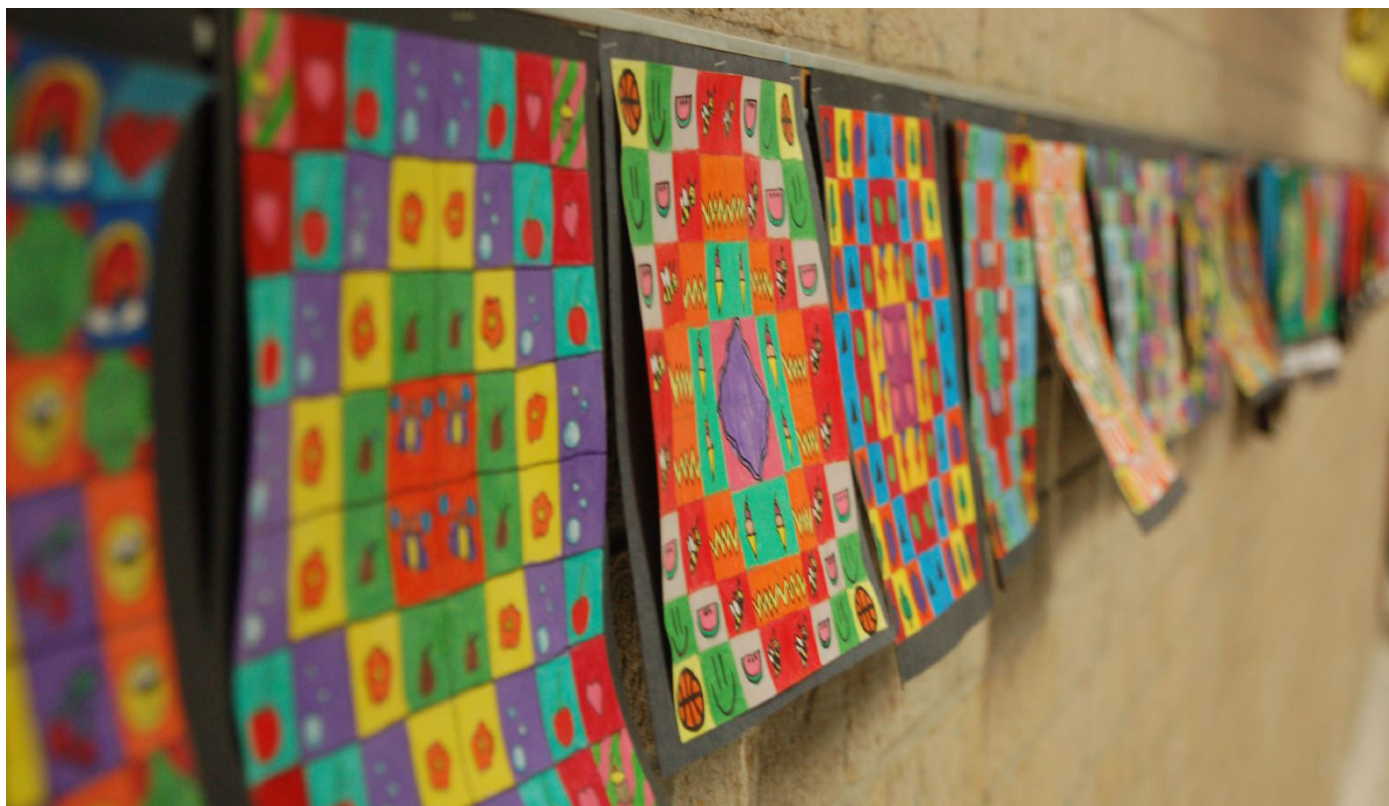
served communities and subgroups (Catterall, et al., 2012; Kisida et al., 2014; Podlozny, 2000). Therefore, we test the hypothesis that students who likely have less access to arts learning opportunities will experience greater treatment impacts. To investigate this possibility, we run separate analyses restricted by school-grade levels served (i.e., elementary and middle), gender, race/ethnicity, FRL and higher poverty status, LEP status, and GT status.

Results

Results, by outcome, for the overall as well as subgroup samples, are provided in tables 3 and 4. We find three statistically significant positive results for the full sample. Increasing students' arts educational experiences reduces the proportion of students receiving a disciplinary infraction by 3.6 percentage points; increases writing achievement by 0.13 of a standard deviation; and increases students' compassion for others by 0.08 of a standard deviation. Estimates are typically in the positive direction for all other outcomes, but fail to achieve traditional levels of statistical significance.

We examine moderating effects by investigating student subgroup impacts. We find numerous positive impacts, particularly on survey measures, with elementary level, LEP, and GT student subgroups. In addition to these

subgroups exhibiting more-pronounced effects on writing achievement and compassion for others, these students demonstrate positive treatment effects on school engagement, college aspirations, arts-facilitated empathy, disposition for arts transfer, and perceived value of the arts. Positive effects on school engagement are particularly noteworthy, as it is the most common positive subgroup finding that was not significant for the overall sample. There is also some evidence to suggest negative impacts with middle school students in terms of school engagement and college aspiration.



Texas.713

Robustness Checks with Survey Data

The availability of survey baseline measures for students in the second cohort allows us to estimate first-year impacts in terms of growth, albeit with a smaller sample of only cohort two students. Nevertheless, this approach serves as a strong robustness check of our estimates. The compassion for others outcome remains positive in this analysis, and it is worth noting that this impact increases by roughly 40 percent, suggesting that these effects may be primarily attributable to impacts that occur in the first year of implementation. This analysis also reveals an overall statistically significant, 0.16 of a standard deviation increase in school engagement and a 5.7 percentage point increase in students' college aspirations.

Survey outcome measures were constructed such that items were standardized and equally weighted. This approach potentially over-weights item responses that are not as strongly correlated to the other items in the construction of the outcome measure. Therefore, we have also examined the sensitivity of our survey results by conducting an analysis where the outcome measures were reconstructed based on factor loadings. This alternative construction of the outcome measures has no impact on the qualitative interpretations of the treatment effect estimates. The comprehensive factor analysis estimates, by outcomes and subgroups, are provided in the appendices.

There also remains the possibility that nonresponse bias could have influenced survey-derived effect estimates. Students in the control group had a significantly higher response rate (9 percentage points) than those in the treatment group. Results from our primary analysis rely on the assumption that survey responses are missing at random and, therefore, not biasing our treatment effect estimates. Relative to the main analytical sample, we do not observe significant differences in surveyed treatment and control group students. While we cannot entirely dismiss concerns of nonresponse bias, we further investigate the sensitivity of these results by conducting an inverse-prob-

ability-weighted regression adjustment (IPWRA) and bounding our treatment effect estimates, such that, on average, we trim 9 percent of the control group survey responses to estimate the bounds of the treatment effect for those students who would have been surveyed regardless of treatment status (see Lee, 2009). Rather than trimming the top and bottom 9 percent of control students, we determined trimmed proportions by response rate disparities within, rather than across the matched pairs. We find that our compassion for others outcome remains positive and statistically significant across specifications with the exception of the lower Lee bounds, which remains in the positive direction but fails to achieve statistical significance. The school engagement outcome is positive for the IPWRA and upper bound specifications, but null with the lower bound specifications. The results from these robustness checks are provided in Table 5.

Discussion & Conclusion

Our investigation is the first large-scale randomized control trial of an arts education program implemented in an authentic educational setting. This program was implemented in a diverse array of elementary and middle schools in the nation's 7th largest school district. We find that increases in students' arts learning experiences significantly improve educational outcomes. Fostering and supporting these experiences lead to improvements in student discipline, writing achievement, and compassion for others. These results are robust and support hypotheses and prior findings that the arts can play a critical role in positively affecting meaningful educational outcomes (Catterall et al., 2012; Deasy, 2002; DiMaggio, 1982; Fiske, 1999; Ruppert, 2006).

In general, the STAAR writing test assesses students' knowledge of mechanics through two components—multiple choice items and composition skills through an open-response expository essay. When we disaggregate this assessment to examine whether overall writing score gains are disproportionately driven by increases in students' mechanical or written composition achievement, we find that the AAI significantly increases student scores on both sections, but effects were greater in magnitude with written composition. Specifically, scores on the writing mechanics section increased by 0.08 of a standard deviation ($p = 0.03$), and scores on the expository essay increased by 0.18 of a standard deviation ($p < 0.01$). This increase in writing achievement, particularly with expository writing composition, is especially noteworthy in that there is limited rigorous evidence for transfer in terms of arts educational interventions bolstering outcomes typically tied to other school subject areas (Winner et al., 2013). The benefits of arts education with writing achievement potentially corroborates and further informs past research showing the benefits of the arts with regard to critical thinking, as assessed through students' written responses and interpretations of original works

of art after an in-depth art museum-facilitated lesson (Bowen et al., 2014).

The main findings from this study were generally robust across student subgroups; however, there were notable variations in subgroup outcomes that could inform policy decisions and future research efforts. We find that reductions in disciplinary infractions appear to be more pronounced with middle school-level, male, African-American, non-FRL, and higher poverty-household students. Recent studies have shown that, with the exception of non-FRL, these students disproportionately incur disciplinary infractions (Anderson & Ritter, 2017; Skiba et al., 2014). Variations in subgroup outcomes are potentially attributable to ceiling effects with groups that have been less likely to incur disciplinary infractions. This consideration appears to be relevant to our sample. Ten percent of the students in our sample had received a disciplinary infraction the previous year; however, these infractions were disproportionately more likely to be incurred by males, middle schoolers, higher poverty students, and African-American students.

We also find that some outcomes are more likely to be statistically significant and positive, as well as larger in magnitude for elementary-level students. In addition to these subgroups exhibiting more-pronounced effects on writing achievement and compassion for others, these students demonstrate positive treatment effects on school engagement, college aspirations, and arts-facilitated empathy. The AAI disproportionately served elementary schools in its first two years, comprising 86 percent of our sample. Program delivery at this school level was the primary focus at this stage of the Initiative, and it is likely the disproportionate attention given to serving these schools may have better ensured fidelity of implementation. On the other hand, the effects unique to middle school students should be interpreted with caution. There were only six middle schools in this study, a circumstance which

lends itself to spurious results (Button et al., 2013). Finally, it is important to take into consideration that educational interventions typically have greater effects with younger students (Heckman, 2006), meaning that even with a larger middle school sample, we might not expect effects of the same magnitude.

Positive effects were also more pronounced with LEP and GT students. LEP and GT students demonstrated numerous positive effects on survey constructs, including school engagement, arts-facilitated empathy, arts transfer disposition, and valuing art. LEP students additionally showed positive effects on college aspirations and desire to consume arts. GT students had an additional positive effect on tolerance. A possible explanation for the strong, positive effects with LEP and GT students is that test-based accountability pressures have yielded circumstances that have withered the scope of K-12 educational opportunities. The majority of schools participating in the AAI had recently experienced pressure with test-based accountability sanctions. These sanctions tend to narrow educational offerings while intensifying the emphasis on student standardized test achievement, especially with those students scoring near critical accountability thresholds (Booher-Jennings, 2005; Deming, Cohodes, Jennings, & Jencks, 2016; Neal, 2010; Neal & Schanzenbach, 2010; West, 2007). The reinjection of the arts in these schools likely expanded enriching offerings and opportunities that plausibly improve student engagement, particularly with LEP and GT students who were more likely to have been adversely affected by school responses to accountability pressures.

There are a few limitations to this study. While random assignment to the AAI treatment allows us to confidently infer a causal relationship between the treatment and assessed outcomes, we remain less certain about the extent to which these results would be achieved in other contexts. In order to participate in the AAI, principals had to have the commitment and desire to improve their schools' arts educational offerings. Our research design ensures that this commitment and desire is balanced across the treatment and control groups. However, such results may not be achievable with schools that lack leaders who are as dedicated to providing and supporting the arts on their campuses. Another context-specific consideration is that Houston is a very arts-rich metropolis with a wealth of cultural institutions and organizations strongly committed to partnering with schools in the community. Smaller or less-arts-rich cities, as well as suburban and rural communities, may lack the resources necessary for providing effective

school-community arts partnerships. However, it is worth noting that there are widespread instances of school-community partnerships with arts organizations across communities of various sizes (Bowen & Kisida, 2017).

Another limitation is that, when defining the treatment for this study, we are restricted to the provision of matching funds and other AAI supports that foster, facilitate, and deliver school-community arts partnerships. We can state with confidence that our findings are a result of increasing arts educational experiences through these partnerships. However, we do not know whether (and which) particular mediators were responsible for these outcomes. The arts educational experiences that students received were incredibly diverse in terms of program format and arts discipline. The AAI director and staff worked with principals to develop plans and program selections that would align with the goals they had for their schools, which typically resulted in students receiving dance, music, theatre, and visual arts educational experiences through each of the various program formats: field trips, in-school professional artist performances, teaching-artist residencies, and programs that took place outside of regular school hours. Due to the variation and endogeneity in program selection across schools, we do not know which particular types of offerings were more likely to bring about desired effects. Finally, these analyses are restricted to shorter-term outcomes, and it remains to be seen if these effects will ultimately compound, serve as a one-time boost, or diminish over time. Critical next steps in this field of study will be to examine whether particular formats and varieties of arts educational programs are more successful in generating desired effects over longer periods of time.

The results of this study provide critical evidence that increasing students' arts educational opportunities has positive impacts on meaningful outcomes. The narrowing of educational offerings and objectives to align with accountability assessments has had adverse effects on the arts in K-12 education. Despite the logic behind narrowing educational offerings, substantial influxes of arts educational experiences do not appear to be detrimental to student growth in outcomes tied to accountability assessments. However, we do find evidence that these reductions pose significant costs. Arts learning experiences benefit students in terms of social, emotional, and academic outcomes. Education policymakers should be mindful and considerate of these benefits when assessing the opportunity costs that come with decisions pertaining to the provision of the arts in schools.

References

- Anderson, K. P., & Ritter, G. W. (2017). Disparate use of exclusionary discipline: Evidence on inequities in school discipline from a U.S. state. *Education Policy Analysis Archives*, 25(49), 1-32.
- Bassok, D., Latham, S., & Rorem, A. (2016). Is kindergarten the new first grade? *AERA Open*, 1(4), 1-31.
- Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. *American Educational Research Journal*, 42(2), 231-268.
- Bowen, D. H., Greene, J. P., and Kisida, B. (2014). Learning to think critically: A visual art experiment. *Educational Researcher*, 42(1), 37-44.
- Bowen, D. H., & Kisida, B. (2017). The art of partnerships: Community resources for arts education. *Phi Delta Kappan*, 98(7), 8-14.
- Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(5), 365.
- Catterall, J. S., Dumais, S. A., & Hampden-Thompson, G. (2012). The arts and achievement in at-risk youth: Findings from four longitudinal studies. Washington, DC: National Endowment for the Arts.
- Catterall, J., & Waldorf, L. (1999). Chicago Arts Partnerships in Education summary evaluation. In E. Fiske (Ed.), *Champions of change: The impact of the arts on learning* (pp. 47-62). Washington, DC: The Arts Education Partnership and The President's Committee on the Arts and the Humanities.
- Deasy, R. (2002). *Critical links: Learning in the arts and student academic and social development*. Washington, DC: Arts Education Partnership.
- Deming, D. J., Cohodes, S., Jennings, J., & Jencks, C. (2016). School accountability, postsecondary attainment, and earnings. *Review of Economics and Statistics*, 98(5), 848-862.
- Dewey, J. (1919). Imagination and expression. *Teachers College Bulletin*, 10(10), 7-15.
- DiMaggio, P. (1982). Cultural capital and school success: The impact of status culture participation on the grades of US high school students. *American Sociological Review*, 47(2), 189-201.
- Eisner, E. (2002). *The arts and the creation of mind*. New Haven, CT: Yale University Press.
- Elpus, K. (2013). Is it the music or is it selection bias? A nationwide analysis of music and non-music students' SAT scores. *Journal of Research in Music Education*, 61(2), 175-194.
- Fiske, E. (Ed.) (1999). *Champions of change: The impact of the arts on learning*. Washington, DC: The Arts Education Partnership and the President's Committee on the Arts and Humanities.
- Gadsden, V. L. (2008). The arts and education: Knowledge generation, pedagogy, and the discourse of learning. *Review of Research in Education*, 32(1), 29-61.
- Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. New York, NY: WW. Norton & Company.
- Gibson, C. (2016). Dancing to the top: How collective action revitalized arts education in Boston. Boston, MA: EdVestors & Boston Public Schools Arts Expansion Partners. Retrieved from www.edvestors.org/wp-content/uploads/2016/05/BPS-ArtsExpansion-Case-Study.pdf
- Government Accountability Office. (2009). *Access to arts education: Inclusion of additional questions in education's planned research would help explain why instruction time has decreased for some students*. Washington, DC: GAO-09-286.
- Greene, J. P., Erickson, H. H., Watson, A. R., & Beck, M. I. (2018). The play's the thing: Experimentally examining the social and cognitive effects of school field trips to live theater performances. *Educational Researcher*, 47(4), 246-254.
- Greene, J. P., Kisida, B., & Bowen, D. H. (2014). The educational value of field trips. *Education Next*, 14(1), 78-86.
- Gromko, J. E. (2005). The effect of music instruction on phonemic awareness in beginning readers. *Journal of Research in Music Education*, 53(3), 199-209.
- Gunzenhauser, M. G. (2003). High-stakes testing and the default philosophy of education. *Theory into Practice*, 42(1), 51-58.
- Hamilton, L. S., Stecher, B. M., Marsh, J. A., McCombs, J. S., & Robyn, A., Russell, J. L., Naftel, S., & Barney, H. (2007). *Standards-based accountability under No Child Left Behind: Experiences of teachers and administrators in three states*. Santa Monica, CA: Rand Corporation.
- Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782), 1900-1902.
- Houston Independent School District (2017). *District and school profiles 2016-17*. Retrieved from https://www.houstonisd.org/cms/lib2/TX01001591/Centricity/Domain/8269/districtdataanalysis/schoolprofiles/2016-2017/Complete%201617%20Profiles_102017.pdf
- Kinney, D. W., & Forsythe, J. L. (2005). The effects of the arts IMPACT curriculum upon student performance on the Ohio fourth-grade proficiency test. *Bulletin of the council for research in music education*, 35-48.

- Kisida, B., Bowen, D. H., & Greene, J. P. (2016). Measuring critical thinking: Results from an art museum field trip experiment. *Journal of Research on Educational Effectiveness*, 9(51), 171-187.
- Kisida, B., Greene, J. P., & Bowen, D. H. (2014). Creating cultural consumers: The dynamics of cultural capital acquisition. *Sociology of Education*, 87(4), 281-295.
- Klineberg, S. L., Wu, J., & Aldape, C. L. (2012). *The Houston arts survey: Participation, perceptions, and prospects*. Kinder Institute for Urban Research. Retrieved from <https://kinder.rice.edu/sites/g/files/bxs1676/f/documents/Kinder%20Houston%20Arts%20Survey%202012.pdf>
- Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, 76, 1071-1102.
- Lee, I. W. and Kim, S. J. (2006). The effects of integrated activity with music and mathematics on musical ability and the mathematical concepts of preschoolers. *The Journal of Korea Open Association for Early Childhood Education*, 11(2), 305-329.
- Ludwig, M. J., Boyle, A., & Lindsay, J. (2017). Review of evidence: Arts integration research through the lens of the Every Student Succeeds Act. Retrieved from <https://www.wallacefoundation.org/knowledge-center/Documents/Arts-Integration-Research-Every-Student-Succeeds-Act-ESSA.pdf>
- Meyer, D., Princiotta, D., & Lanahan, L. (2004). *The Summer After Kindergarten: Children's Activities and Library Use by Household Socioeconomic Status* (NCES 2004-037). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science*, 22(11), 1425-1433.
- Moreno S., Marques, C., Santos, A. Santos, M., Castro, S. L., & Besson, M. (2009). Musical training influences linguistic abilities in 8-year-old children: More evidence for brain plasticity. *Cerebral Cortex*, 19(3), 712-723.
- Neal, D. (2010). Aiming for efficiency rather than proficiency. *Journal of Economic Perspectives*, 24(3), 119-132.
- Neal, D., & Schanzenbach, D. W. (2010). Left behind by design: Proficiency counts and test-based accountability. *The Review of Economics and Statistics*, 92(2), 263-283.
- Neville, H. (2008). Effects of music training on brain and cognitive development in under-privileged 3- to 5-year-old children: Preliminary results. In B. Rich and C. Asbury (eds.), *Learning, Arts, and the Brain: The Dana Consortium Report on Arts and Cognition* (pp. 105-106). New York, NY: The Dana Foundation.
- Perille (2016, November 29). A blueprint for successful arts education: Five strategies for investing in students' creative capital. *Education Week*. Retrieved from <https://www.edweek.org/ew/articles/2016/11/30/a-blueprint-for-successful-arts-education.html>
- Podlozny, A. (2000). Strengthening verbal skills through the use of classroom drama: A clear link. *Journal of Aesthetic Education*, 34(3-4), 91-104.
- Rabkin, N., & Hedberg, E. C. (2011). Arts education in America: What the declines mean for arts participation. Based on the 2008 Survey of Public Participation in the Arts. Research Report# 52. *National Endowment for the Arts*.
- Redford, J., Burns, S., & Hall, L. J. (2018). *The Summer After Kindergarten: Children's Experiences by Socioeconomic Characteristics* (NCES 2018-160). U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Rowe, M., Werber, L., Kaganoff, T., & Robyn, A. (2004). *Arts education partnerships*. Santa Monica, CA: RAND Corporation.
- Ruppert, S. S. (2006). *Critical evidence: How the arts benefit student achievement*. Washington, DC: National Assembly of State Arts Agencies.
- Schellenberg, E. G. (2004). Music lessons enhance IQ. *Psychological Science*, 15(8), 511-514.
- Silk, Y. (2016). *State of the arts in Chicago Public Schools: Progress report 2015-16*. Retrieved from www.joomag.com/magazine/ingenuity-state-of-the-arts-progress-report2015-2016/0495510001479939351?short.
- Skiba, R. J., Chung, C., Trachok, M., Baker, T., Sheya, A., & Hughes, R. (2014). Parsing disciplinary disproportionality: Contributions of infraction, student, and school characteristics to out-of-school suspension and expulsion. *American Educational Research Journal*, 51(4), 640-670.
- Standley, J. M. (2008). Does music instruction help children learn to read? Evidence of a meta-analysis. *Update: Applications of Research in Music Education*, 27(1), 17-32.
- Thomas, M. K., Singh, P., & Klopfenstein, K. (2015). Arts education and the high school dropout problem. *Journal of Cultural Economics*, 39(4), 327-339.
- von Zastrow, C. E., & Janc, H. (2004). *Academic atrophy: The condition of the liberal arts in America's public schools*. Council for Basic Education.
- West, M. (2007). Testing, learning, and teaching: The effects of test-based accountability on student achievement and instructional time in core academic subjects. In Chester E. Finn, Jr. & Diane Ravitch (Eds.) *Beyond the Basics: Achieving a Liberal Education for All Children* (pp. 45-61). Washington, DC: The Fordham Institute.
- Winner, E., & Cooper, M. (2000). Mute those claims: No evidence (yet) for a causal link between arts study and academic achievement. *Journal of Aesthetic Education*, 34(3/4), 11-75.
- Winner, E., & Goldstein, T. R. Vincent-Lancrin, S. (2013). *Art for art's sake? The impacts of arts education*. Paris, France: OECD Publishing.
- Winner, E., & Hetland, L. (2001). The arts and academic achievement: What the evidence shows, Executive summary. *The Journal of Aesthetic Education*, 34, 3-4.
- Yee, V. (2014, April 7). Arts education lacking in low-income areas of New York City, report says. *The New York Times*. Retrieved from <https://www.nytimes.com/2014/04/07/nyregion/arts-education-lacking-in-low-income-areas-of-new-york-city-report-says.html>
- Young Audiences of Houston (2014). Arts Access Initiative: 2013-14 summary.
- Zimmerman, E. (2009). Reconceptualizing the role of creativity in art education theory and practice. *Studies in Art Education*, 50(4), 382-399.

Table 1

Survey Outcomes

Outcome	Items	Source(s)	Cronbach's Alpha
School Engagement	<ul style="list-style-type: none"> School work is interesting. School work makes me think about things in new ways. School work is not very enjoyable. This school is a happy place for me to be. This school offers lots of different types of programs, classes, and activities to keep me interested in school. 	Tripod Student Perceptions; Austin ISD Climate Survey	0.73
College Aspiration	<ul style="list-style-type: none"> I plan to go to college. 	Austin ISD Climate Survey	N/A
Arts-Facilitated Empathy	<ul style="list-style-type: none"> Works of art, like paintings, music, dance performances, and plays, help me understand what life was like in another time or place. I can learn about my classmates by listening to them talk about works of art. Works of art help me imagine what life is like for someone else. 	Greene, Kisida, & Bowen (2014)	0.67
Compassion for Others	<ul style="list-style-type: none"> I want to help people who get treated badly. I am not really interested in how other people feel. 	Greene, Kisida, & Bowen (2014)	0.47
Tolerance	<ul style="list-style-type: none"> I think people can have different opinions about the same thing. I appreciate hearing views different from my own. People who disagree with my point of view bother me. 	Greene, Kisida, & Bowen (2014)	0.38
Cultural Consumption	<ul style="list-style-type: none"> I plan to go to art museums and galleries when I am an adult. I plan to go to the theater for plays, musicals, and performances when I am an adult. I plan to go to music concerts when I am an adult. I plan to go to dance performances when I am an adult. 	Greene, Kisida, & Bowen (2014)	0.73
Arts Transfer Disposition	<ul style="list-style-type: none"> The arts can help me be a better student. Things I have learned in school can help me understand the arts. Learning about the arts can help me understand things I learn in school. 	NAEA-AAMD Art Museum Impact Survey	0.78
Value the Arts	<ul style="list-style-type: none"> The arts are interesting to me. The arts are an important part of my life. I enjoy talking about the arts. I think artists do important work. Art is one of my favorite subjects in school. I would like to take more classes in the arts when I get to high school and college. 	NAEA-AAMD Art Museum Impact Survey	0.85

Note: Due to there being only two items that comprised Compassion for Others and low internal consistency with Tolerance, we have analyzed these items, individually, that were used to form this construct. We have run results with these outcomes both as a single construct and by examining each of the survey items that comprise the construct individually. We primarily report analyses with these outcomes as single constructs; however, we note any instances in the ITT overall and subgroup analyses where individual items produce results that substantially deviate from those of their constructs.

Table 2

Descriptive Statistics by Treatment Status

Variable	Treatment	Control	Difference
Grade Level	5.594 (0.287)	5.701 (0.274)	-0.107 (0.396)
Female	0.486 (0.006)	0.489 (0.008)	-0.003 (0.010)
FRL Status:			
Not FRL	0.136 (0.023)	0.134 (0.020)	0.002 (0.031)
Free Lunch Eligible	0.301 (0.017)	0.289 (0.021)	0.013 (0.027)
Reduced Lunch Eligible	0.074 (0.007)	0.070 (0.007)	0.004 (0.010)
Poverty	0.489 (0.030)	0.508 (0.022)	-0.019 (0.037)
Race/Ethnicity:			
African-American	0.232 (0.043)	0.243 (0.045)	-0.010 (0.062)
Hispanic/Latinx	0.708 (0.046)	0.701 (0.044)	0.007 (0.064)
White	0.027 (0.010)	0.029 (0.012)	-0.003 (0.015)
Special Education (SPED)	0.060 (0.005)	0.068 (0.009)	-0.008 (0.009)
Limited English Proficiency (LEP)	0.302 (0.038)	0.283 (0.030)	0.019 (0.048)
Baseline Reading (standardized)	0.054 (0.071)	-0.128 (0.048)	0.182* (0.085)
Baseline Math (standardized)	0.018 (0.067)	-0.076 (0.057)	0.094 (0.088)
Prior Year Disciplinary Infraction	0.081 (0.019)	0.123 (0.027)	-0.042 (0.033)
Baseline Absences	4.695 (0.318)	5.389 (0.397)	-0.694 (0.507)
Sample Size	5,333	5,215	10,548

Note: ** statistically significant (two-tailed) at $p < 0.01$; * significant at $p < 0.05$; standard errors in parentheses adjusted for school-grade clustering. When “baseline” is specified for time-sensitive academic achievement, it is for 2014–15 for the first cohort and 2015–16 for the second cohort. Test scores standardized relative to the rest of HISD with a mean of zero and standard deviation of one.

Table 3

Administrative Data-Derived Outcomes by Subgroup

Population	N	Discipline	Absences	Math	Reading	Science	Writing
Overall	4,063–10,548	-0.036* (0.015)	0.061 (0.128)	0.014 (0.063)	-0.019 (0.019)	-0.046 (0.065)	0.127** (0.046)
Elementary	2,547–5,565	-0.002 (0.007)	0.192 (0.136)	0.013 (0.042)	-0.029 (0.032)	0.009 (0.049)	0.179** (0.065)
Middle	1,516–4,983	-0.073* (0.028)	-0.087 (0.209)	-0.006 (0.121)	-0.009 (0.021)	-0.144 (0.132)	0.030 (0.023)
Female	2,009–5,140	-0.020 (0.011)	0.238 (0.142)	0.019 (0.070)	-0.021 (0.022)	-0.022 (0.067)	0.136** (0.047)
Male	2,054–5,408	-0.053** (0.020)	-0.100 (0.177)	0.010 (0.057)	-0.019 (0.021)	-0.062 (0.066)	0.106* (0.050)
Af.-Amer.	950–2,503	-0.049** (0.015)	-0.275 (0.304)	0.058 (0.045)	-0.062 (0.035)	0.061 (0.059)	0.031 (0.063)
Hisp-Latx.	2,856–7,436	-0.035 (0.018)	0.130 (0.147)	-0.016 (0.073)	-0.012 (0.022)	-0.064 (0.073)	0.134* (0.057)
Not FRL	602–1,421	-0.059** (0.022)	0.562* (0.221)	0.086 (0.082)	0.050 (0.044)	-0.083 (0.135)	0.225 (0.116)
FRL	1,485–3,875	-0.027* (0.013)	0.112 (0.117)	-0.026 (0.087)	-0.031 (0.020)	-0.073 (0.074)	0.124* (0.057)
Poverty	1,975–5,255	-0.039* (0.018)	-0.077 (0.173)	0.020 (0.045)	-0.031 (0.025)	-0.007 (0.050)	0.094* (0.042)
LEP	1,110–3,089	-0.023 (0.017)	-0.079 (0.185)	-0.030 (0.056)	-0.025 (0.031)	0.007 (0.062)	0.266* (0.100)
GT	731–1,927	-0.014 (0.010)	0.127 (0.149)	0.210 (0.166)	0.047 (0.035)	0.005 (0.150)	0.175 (0.094)

Note: ** statistically significant (two-tailed) at $p < 0.01$; * significant at $p < 0.05$; standard errors in parentheses adjusted for school-grade clustering. Test scores are standardized with a mean of zero and a standard deviation of one.

Table 4

Survey Data-Derived Outcomes by Subgroup

Population	N	School Engagement	College Aspiration	Arts Empathy	Compassion	Tolerance	Cultural Consumption	Arts Transfer Disposition	Values Art
Overall	6,241–6,325	0.093 (0.051)	0.015 (0.018)	0.039 (0.036)	0.080* (0.032)	-0.004 (0.036)	0.029 (0.036)	0.054 (0.038)	0.060 (0.042)
Elementary	3,613–3,660	0.260** (0.056)	0.069** (0.021)	0.097* (0.047)	0.150** (0.038)	0.041 (0.040)	0.074 (0.050)	0.079 (0.047)	0.054 (0.055)
Middle	2,628–2,665	-0.136* (0.047)	-0.053* (0.018)	-0.021 (0.045)	0.010 (0.048)	-0.050 (0.064)	-0.005 (0.040)	0.030 (0.057)	0.091 (0.064)
Female	3,152–3,185	0.120* (0.058)	0.020 (0.020)	0.069 (0.044)	0.059 (0.040)	-0.012 (0.043)	0.051 (0.041)	0.048 (0.051)	0.042 (0.049)
Male	3,093–3,140	0.067 (0.053)	0.008 (0.023)	-0.004 (0.042)	0.094* (0.042)	-0.003 (0.041)	-0.004 (0.049)	0.054 (0.039)	0.068 (0.046)
Af.-Amer.	1,444–1,472	0.146* (0.070)	-0.040 (0.023)	-0.047 (0.059)	0.046 (0.041)	-0.110 (0.068)	0.030 (0.056)	0.069 (0.064)	0.032 (0.066)
Hisp-Latx.	4,333–4,386	0.065 (0.055)	0.042 (0.021)	0.049 (0.038)	0.089* (0.043)	0.010 (0.039)	0.027 (0.043)	0.025 (0.035)	0.046 (0.038)
Not FRL	840–848	0.162* (0.073)	-0.030 (0.036)	0.078 (0.114)	0.091 (0.063)	-0.012 (0.067)	0.033 (0.079)	0.114 (0.102)	0.207 (0.115)
FRL	2,345–2,381	0.077 (0.060)	0.025 (0.025)	0.064 (0.044)	0.092 (0.048)	-0.007 (0.046)	0.042 (0.044)	0.070 (0.049)	0.091 (0.055)
Poverty	3,058–3,098	0.086 (0.056)	0.022 (0.018)	-0.008 (0.039)	0.062 (0.038)	0.001 (0.044)	0.012 (0.040)	0.009 (0.039)	-0.021 (0.041)
LEP	1,970–1,993	0.207** (0.072)	0.100** (0.024)	0.133* (0.058)	0.222** (0.056)	0.041 (0.048)	0.165* (0.065)	0.137* (0.061)	0.115* (0.057)
GT	1,265–1,281	0.158* (0.073)	0.042 (0.027)	0.201** (0.069)	0.232** (0.046)	0.136** (0.051)	0.092 (0.056)	0.176** (0.065)	0.232** (0.074)

Note: ** statistically significant (two-tailed) at $p < 0.01$; * significant at $p < 0.05$; standard errors in parentheses adjusted for school-grade clustering. When investigating individual items for Compassion for Others and Tolerance, there are some qualitative differences in the interpretations of individual survey item effects relative to those for the constructed outcomes. For the Compassion item “I want to help people who get treated badly,” there is a significant positive effect with females ($p = 0.03$) and FRL-eligible students ($p = 0.03$), and a lack of a significant effect with males ($p = 0.08$). For the Compassion item “I am not really interested in how other people feel,” there is a lack of a significant effect with the overall sample ($p = 0.09$) and Hispanic/Latinx students ($p = 0.09$). For the Tolerance item “I think people can have different opinions about the same thing,” there is a significant negative effect with African-Americans ($p = 0.01$). For the Tolerance item “I appreciate hearing views different from my own,” there is a significant positive effect with LEP students ($p = 0.03$). For the Tolerance item “People who disagree with my point of view bother me,” there is a lack of a significant effect with GT students ($p = 0.12$).

Table 5

Robustness Checks for Survey Outcomes

Outcome	N	ITT	2SLS	Pre-Post Cohort 2	Factor Analysis	IPWRA ATE	Lee Bounds	
							Upper	Lower
School Engagement	6,325	0.093 (0.052)	0.095 (0.051)	0.155** (0.048)	0.080 (0.043)	0.074** (0.025)	0.164** (0.055)	0.045 (0.073)
College Aspiration	6,311	0.015 (0.018)	0.015 (0.018)	0.057* (0.023)	N/A	0.010 (0.012)	0.080** (0.029)	-0.035 (0.032)
Arts-Facilitated Empathy	6,249	0.039 (0.036)	0.039 (0.036)	0.051 (0.047)	0.029 (0.028)	0.024 (0.027)	0.118* (0.046)	-0.020 (0.066)
Compassion for Others	6,290	0.080* (0.032)	0.081* (0.033)	0.113** (0.036)	0.047* (0.019)	0.071** (0.026)	0.183** (0.050)	0.033 (0.056)
Tolerance	6,299	-0.004 (0.036)	-0.004 (0.036)	0.006 (0.043)	-0.001 (0.021)	-0.016 (0.026)	0.057 (0.047)	-0.060 (0.064)
Cultural Consumption	6,241	0.030 (0.036)	0.030 (0.036)	0.031 (0.047)	0.022 (0.030)	0.030 (0.025)	0.086 (0.050)	-0.035 (0.062)
Arts Transfer Disposition	6,259	0.054 (0.038)	0.055 (0.038)	0.042 (0.051)	0.044 (0.032)	0.040 (0.027)	0.133** (0.042)	-0.011 (0.072)
Value the Arts	6,285	0.060 (0.043)	0.061 (0.043)	0.016 (0.045)	0.036 (0.028)	0.035 (0.027)	0.129** (0.043)	0.003 (0.073)

Note: ** statistically significant (two-tailed) at $p < 0.01$; * significant at $p < 0.05$; standard errors in parentheses adjusted for school-grade clustering. Pre-post analyses are for cohort two students only; this restriction reduces the sample sizes to a range of 3,511 to 3,625 observations. Lee Bounds “snip” the upper and lower tails of the dependent variable of interest distribution with the control group, which had a higher response rate than the treatment group; the upper and lower bound snips reduced the sample sizes to a range of 5,381 to 5,700 observations. College aspiration is not included because results are based on a single survey item.

Appendices

Survey Participant Demographics and Treatment-Control Comparisons

Variable	Survey Participants	Survey Population	Part.-Pop. Difference	Survey Treat	Survey Control	T-C Difference
Grade Level	5.674 (0.232)	5.958 (0.253)	-0.284 (0.184)	5.693 (0.346)	5.656 (0.314)	0.037 (0.466)
Female	0.504 (0.007)	0.476 (0.011)	0.027* (0.014)	0.502 (0.010)	0.505 (0.011)	-0.003 (0.015)
FRL Status:						
Not FRL	0.134 (0.021)	0.132 (0.017)	0.002 (0.018)	0.139 (0.034)	0.130 (0.026)	0.009 (0.043)
Free Lunch	0.302 (0.017)	0.290 (0.020)	0.012 (0.014)	0.305 (0.024)	0.299 (0.024)	0.007 (0.034)
Reduced Lunch	0.074 (0.006)	0.068 (0.008)	0.006 (0.006)	0.076 (0.008)	0.073 (0.008)	0.004 (0.011)
Poverty	0.490 (0.022)	0.509 (0.022)	-0.019 (0.016)	0.480 (0.035)	0.499 (0.026)	-0.019 (0.044)
Race/Ethnicity:						
African-American	0.233 (0.032)	0.255 (0.047)	-0.022 (0.037)	0.244 (0.052)	0.222 (0.039)	0.022 (0.065)
Hispanic/Latinx	0.694 (0.034)	0.709 (0.048)	-0.015 (0.039)	0.679 (0.058)	0.707 (0.039)	-0.029 (0.070)
White	0.034 (0.011)	0.017 (0.004)	0.018* (0.009)	0.032 (0.016)	0.037 (0.016)	-0.005 (0.023)
SPED	0.048 (0.004)	0.097 (0.010)	-0.049** (0.010)	0.045 (0.006)	0.050 (0.005)	-0.006 (0.008)
LEP	0.315 (0.029)	0.318 (0.038)	-0.003 (0.037)	0.311 (0.047)	0.319 (0.035)	-0.007 (0.058)
Baseline Reading	0.046 (0.051)	-0.151 (0.072)	0.197** (0.067)	0.157 (0.085)	-0.055 (0.048)	0.212* (0.097)
Baseline Math	0.068 (0.054)	-0.122 (0.061)	0.190** (0.058)	0.140 (0.087)	0.002 (0.062)	0.138 (0.106)
Baseline Absences	4.289 (0.200)	7.311 (0.571)	-2.344** (0.421)	4.705 (0.368)	5.207 (0.363)	-0.502 (0.514)
Baseline Discipline	0.069 (0.014)	0.122 (0.027)	-0.052* (0.020)	0.062 (0.020)	0.075 (0.019)	-0.013 (0.027)
Sample Size	6,340	8,437	8,437	3,026	3,314	6,340

Note: ** statistically significant (two-tailed) at $p < 0.01$; * significant at $p < 0.05$; standard errors adjusted for school-grade clustering. Survey participation column is for the spring administration and is relative to the targeted survey population. When “baseline” is specified for time-sensitive academic achievement, it is for 2014–15 for the first cohort and 2015–16 for the second cohort. Test scores standardized relative to the rest of HISD with a mean of zero and standard deviation of one.

First Year Administrative and Survey Results—ITT and ATE

Outcome	N	ITT (s.e.)	1 st Stage (s.e.)	2SLS (s.e.)
Disciplinary Infraction	10,548	-0.036* (0.015)	0.961** (0.003)	-0.038* (0.015)
Absences	10,548	0.061 (0.128)	0.961** (0.003)	0.063 (0.132)
Math (4 th -8 th)	10,130	0.014 (0.063)	0.977** (0.002)	0.014 (0.064)
Reading (4 th -8 th)	10,140	-0.046 (0.065)	0.977** (0.002)	-0.020 (0.020)
Science (5 th and 8 th)	4,063	-0.046 (0.065)	0.975** (0.003)	-0.047 (0.065)
Writing (4 th and 7 th)	4,352	0.127** (0.046)	0.977** (0.003)	0.130** (0.046)
School Engagement	6,325	0.093 (0.051)	0.987** (0.002)	0.095 (0.051)
College Aspiration	6,311	0.015 (0.018)	0.987** (0.002)	0.015 (0.018)
Arts-Facilitated Empathy	6,249	0.039 (0.036)	0.987** (0.002)	0.039 (0.036)
Compassion for Others	6,290	0.080* (0.032)	0.987** (0.002)	0.081* (0.033)
Tolerance	6,299	-0.004 (0.036)	0.987** (0.002)	-0.004 (0.036)
Cultural Learning	6,316	0.037 (0.039)	0.987** (0.002)	0.038 (0.039)
Cultural Consumption	6,241	0.030 (0.036)	0.987** (0.002)	0.030 (0.036)
Arts Transfer Disposition	6,259	0.054 (0.038)	0.987** (0.002)	0.055 (0.038)
Value the Arts	6,285	0.060 (0.042)	0.987** (0.002)	0.061 (0.043)

Note: ** statistically significant (two-tailed) at $p < 0.01$; * significant at $p < 0.05$. Standard errors are in parentheses and have been clustered at the school-grade level. Test scores are standardized with a mean of zero and a standard deviation of one.

Local Average Treatment Effects for Administrative Data-Derived Outcomes

Population	N	Discipline	Absences	Math	Reading	Science	Writing
Overall	4,063–10,548	-0.038* (0.015)	0.063 (0.132)	0.014 (0.064)	-0.020 (0.020)	-0.047 (0.065)	0.130** (0.046)
Elementary	2,547–5,565	-0.002 (0.007)	0.201 (0.141)	0.014 (0.043)	-0.030 (0.032)	0.010 (0.050)	0.184** (0.065)
Middle	1,516–4,983	-0.075** (0.028)	-0.091 (0.211)	-0.006 (0.120)	-0.009 (0.021)	-0.147 (0.124)	0.030 (0.021)
Female	2,009–5,140	-0.021 (0.012)	0.247 (0.146)	0.019 (0.071)	-0.021 (0.022)	-0.023 (0.068)	0.140** (0.047)
Male	2,054–5,408	-0.055** (0.020)	-0.104 (0.183)	0.011 (0.058)	-0.019 (0.021)	-0.064 (0.066)	0.108* (0.050)
Af.-Amer.	950–2,503	-0.053** (0.016)	-0.295 (0.322)	0.061 (0.046)	-0.064 (0.036)	0.064 (0.059)	0.032 (0.064)
Hisp-Latx.	2,856–7,436	-0.036* (0.018)	0.134 (0.151)	-0.016 (0.073)	-0.012 (0.022)	-0.066 (0.073)	0.136* (0.057)
Not FRL	602–1,421	-0.061** (0.022)	0.581* (0.225)	0.088 (0.082)	0.051 (0.044)	-0.084 (0.132)	0.231* (0.114)
FRL	1,485–3,875	-0.028* (0.013)	0.115 (0.119)	-0.027 (0.088)	-0.031 (0.021)	-0.074 (0.073)	0.127* (0.057)
Poverty	1,975–5,255	-0.041* (0.018)	-0.081 (0.181)	0.021 (0.046)	-0.032 (0.026)	-0.007 (0.050)	0.096* (0.042)
LEP	1,110–3,089	-0.023 (0.017)	-0.082 (0.189)	-0.031 (0.057)	-0.026 (0.031)	0.007 (0.061)	0.271** (0.100)
GT	731–1,927	-0.014 (0.010)	0.130 (0.150)	0.213 (0.166)	0.048 (0.035)	0.005 (0.147)	0.178 (0.092)

Note: ** statistically significant (two-tailed) at $p < 0.01$; * significant at $p < 0.05$; standard errors in parentheses adjusted for school-grade clustering. Test scores are standardized with a mean of zero and a standard deviation of one.

Local Average Treatment Effects for Survey Data-Derived Outcomes

Population	N	School Engagement	College Aspiration	Arts Empathy	Compassion	Tolerance	Cultural Consumption	Arts Transfer Disposition	Values Art
Overall	6,241–6,325	0.095 (0.051)	0.015 (0.018)	0.039 (0.036)	0.081* (0.033)	-0.004 (0.036)	0.030 (0.036)	0.055 (0.038)	0.061 (0.043)
Elementary	3,613–3,660	0.266** (0.057)	0.071** (0.021)	0.099* (0.048)	0.153** (0.039)	0.042 (0.041)	0.075 (0.050)	0.080 (0.047)	0.055 (0.055)
Middle	2,628–2,665	-0.136** (0.045)	-0.053** (0.018)	-0.022 (0.044)	0.010 (0.047)	-0.050 (0.062)	-0.005 (0.038)	0.031 (0.055)	0.091 (0.062)
Female	3,152–3,185	0.122* (0.058)	0.020 (0.020)	0.070 (0.044)	0.060 (0.040)	-0.012 (0.043)	0.051 (0.041)	0.049 (0.051)	0.043 (0.049)
Male	3,093–3,140	0.068 (0.053)	0.008 (0.023)	-0.004 (0.042)	0.095* (0.042)	-0.003 (0.041)	-0.004 (0.049)	0.055 (0.039)	0.069 (0.046)
Af.-Amer.	1,444–1,472	0.148* (0.069)	-0.039 (0.023)	-0.048 (0.058)	0.047 (0.041)	-0.112 (0.068)	0.030 (0.055)	0.070 (0.063)	0.032 (0.065)
Hisp.-Latx	4,333–4,386	0.066 (0.055)	0.042* (0.022)	0.050 (0.038)	0.090* (0.043)	0.010 (0.040)	0.028 (0.043)	0.026 (0.035)	0.046 (0.038)
Not FRL	840–848	0.164 (0.072)	-0.031 (0.035)	0.079 (0.113)	0.093 (0.062)	-0.012 (0.067)	0.034 (0.078)	0.116 (0.100)	0.210 (0.114)
FRL	2,345–2,381	0.079 (0.060)	0.025 (0.025)	0.069 (0.045)	0.095* (0.048)	-0.007 (0.046)	0.043 (0.044)	0.074 (0.050)	0.095 (0.055)
Poverty	3,058–3,098	0.087 (0.056)	0.022 (0.018)	-0.008 (0.039)	0.063 (0.038)	0.001 (0.044)	0.012 (0.041)	0.010 (0.039)	-0.022 (0.041)
LEP	1,970–1,993	0.210** (0.072)	0.102** (0.024)	0.135* (0.058)	0.225** (0.056)	0.042 (0.048)	0.167* (0.066)	0.139* (0.061)	0.116* (0.057)
GT	1,265–1,281	0.160* (0.073)	0.043 (0.027)	0.204** (0.069)	0.236** (0.046)	0.138** (0.050)	0.094 (0.055)	0.179** (0.065)	0.235** (0.073)

Note: ** statistically significant (two-tailed) at $p < 0.01$; * significant at $p < 0.05$; standard errors in parentheses adjusted for school-grade clustering.

Survey-Derived Outcomes with Factor Analysis

Population	N	School Engagement	Arts Empathy	Compassion	Tolerance	Cultural Consumption	Arts Transfer Disposition	Values Art
Overall	6,122–6,311	0.080 (0.043)	0.029 (0.028)	0.047* (0.019)	-0.001 (0.021)	0.022 (0.030)	0.044 (0.032)	0.036 (0.028)
Elementary	3,540–3,652	0.224** (0.046)	0.076* (0.037)	0.088** (0.022)	0.027 (0.024)	0.060 (0.042)	0.062 (0.040)	0.044 (0.035)
Middle	2,628–2,665	-0.114** (0.037)	-0.019 (0.034)	0.006 (0.028)	-0.028 (0.039)	-0.008 (0.036)	0.027 (0.049)	0.038 (0.041)
Female	3,096–3,182	0.106* (0.049)	0.051 (0.034)	0.034 (0.022)	-0.007 (0.026)	0.039 (0.035)	0.038 (0.044)	0.014 (0.033)
Male	3,026–3,129	0.055 (0.045)	-0.002 (0.033)	0.057* (0.024)	0.002 (0.024)	-0.004 (0.041)	0.045 (0.033)	0.049 (0.029)
Af.-Amer.	1,408–1,469	0.122* (0.058)	-0.047 (0.045)	0.023 (0.024)	-0.073 (0.039)	0.031 (0.048)	0.055 (0.055)	0.023 (0.043)
Hisp-Latx.	4,254–4,377	0.056 (0.046)	0.041 (0.030)	0.053* (0.024)	0.012 (0.024)	0.013 (0.036)	0.020 (0.030)	0.019 (0.023)
Not FRL	832–845	0.131* (0.063)	0.066 (0.088)	0.050 (0.037)	-0.006 (0.039)	0.046 (0.071)	0.095 (0.086)	0.136 (0.079)
FRL	2,302–2,380	0.067 (0.050)	0.053 (0.034)	0.054 (0.028)	0.002 (0.027)	0.023 (0.037)	0.059 (0.042)	0.036 (0.034)
Poverty	2,988–3,092	0.073 (0.046)	-0.011 (0.030)	0.036 (0.022)	0.001 (0.026)	0.008 (0.034)	0.008 (0.033)	-0.002 (0.027)
LEP	1,925–1,989	0.179** (0.061)	0.111* (0.045)	0.135** (0.031)	0.038 (0.030)	0.122* (0.054)	0.112* (0.051)	0.077* (0.036)
GT	1,250–1,280	0.131* (0.063)	0.154** (0.054)	0.134** (0.026)	0.086** (0.029)	0.067 (0.047)	0.149** (0.055)	0.150** (0.052)

Note: ** statistically significant (two-tailed) at $p < 0.01$; * significant at $p < 0.05$; standard errors in parentheses adjusted for school-grade clustering. College aspiration is not included because results are based on a single survey item.

Mission

The Kinder Institute for Urban Research builds better cities and improves people's lives by bringing together data, research, engagement, and action.

